Université Paris I, Panthéon - Sorbonne

Première Année Master M.A.E.F. 2013 – 2014

Statistiques II

Contrôle continu n°1, février 2014

Examen de 1h30. Tout document ou calculatrice est interdit.

1. **(Sur 9 points)** On suppose que $(\varepsilon_i)_{i\in\mathbf{N}}$ est une suite de v.a.i.i.d. de loi uniforme sur $[0,\theta]$ avec $\theta > 0$ un réel inconnu. On définit la suite $(X_t)_{t\in\mathbf{N}}$ par $X_0 = 0$ et la relation suivante:

$$X_{t+1} = \max\big(X_t\,,\,\varepsilon_{t+1}\big) \qquad \text{pour tout } t \in \mathbf{N}.$$

   (a) Montrer que pour tout $t \in \mathbf{N}^*$, $X_t = \max\big(\varepsilon_1, \cdots, \varepsilon_t\big)$ **(1 pt)**.

   (b) Pour $t \in \mathbf{N}^*$, montrer que la densité de probabilité de X_t par rapport à la mesure de Lebesgue sur $\mathbf{R}$ vaut: $f_t(x) = t\,x^{t-1}\,\theta^{-t}\mathbb{I}_{t\in[0,\theta]}$ **(1.5 pts)**.

   (c) Déterminer la tendance et la saisonnalité de (X_t) **(1 pt)**. La série (X_t) est-elle stationnaire **(0.5 pts)**?

   (d) Montrer que les variables X_t ne sont pas indépendantes **(1.5 pts)**.

   (e) Montrer que $X_n \xrightarrow[n\to+\infty]{\mathcal{P}} \theta$ **(1.5 pts)**.

   (f) Soit $Z_n = n(\theta - X_n)$. Déterminer la densité de probabilité de Z_n **(1 pt)**. En déduire que $Z_n \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{E}(1/\theta)$ où $\mathcal{E}(1/\theta)$ est une loi exponentielle de paramètre $1/\theta$ **(1 pt)**.

Proof. (a) On a par itération

$$X_{t+1} = \max\Big(X_t\,,\,\varepsilon_{t+1}\Big) = \max\Big(\max\Big(X_{t-1}\,,\,\varepsilon_t\Big)\,,\,\varepsilon_{t+1}\Big) = \max\Big(X_{t-1}\,,\,\varepsilon_t\,,\,\varepsilon_{t+1}\Big) = \cdots = \max\Big(X_0\,\varepsilon_1\,,\,\varepsilon_t\,,\,\varepsilon_{t+1}\Big)$$

Avec $X_0 = 0$, et $\varepsilon_i \geq 0$ pour tout i, on en déduit que on en déduit que $X_t = \max\Big(\varepsilon_1, \cdots, \varepsilon_t\Big)$.

(b) Pour $x \in [0,\theta]$, $\mathbb{P}(X_t \leq x) = \mathbb{P}(\varepsilon_1 \leq x, \cdots, \varepsilon_t \leq x) = \mathbb{P}(\varepsilon_1 \leq x) \times \cdots \times \mathbb{P}(\varepsilon_1 \leq x)$ par l'indépendance des ε_i. Ainsi $\mathbb{P}(X_t \leq x) = (x/\theta)^n$. On obtient alors la densité de X_t en dérivant cette fonction de répartition.

(c) On a $\mathbb{E}X_t = \int_0^\theta x f_t(x)dx = t\theta/(t+1)$. (X_t) a donc pour tendance la fonction $t \mapsto t\theta/(t+1)$ et pas de saisonnalité. Comme la tendance n'est pas une constante, (X_t) n'est pas stationnaire.

(d) On a pour $t > s$, $\mathbb{P}(X_t < \theta/2 \text{ et } X_s > \theta/2) = 0$ car $X_t \geq X_s$. Mais $\mathbb{P}(X_t < \theta/2)\mathbb{P}(X_s > \theta/2) \neq 0$, donc il n'y a pas indépendance.

(e) Soit $Y_n = \theta - X_n$ pour $n \in \mathbf{N}$. On a $Y_n \geq 0$ et $\mathbb{E}Y_n = \theta/(n+1)$. Donc $\mathbb{E}Y_n \xrightarrow[n\to\infty]{} 0$. D'après l'inégalité de Markov, pour tout $\varepsilon > 0$, $P(Y_n \geq \varepsilon) \leq \mathbb{E}Y_n/\varepsilon \xrightarrow[n\to\infty]{} 0$: $Y_n \xrightarrow[n\to+\infty]{\mathcal{P}} 0$ et donc $X_n \xrightarrow[n\to+\infty]{\mathcal{P}} \theta$.

(f) Pour tout $n \in \mathbf{N}^*$, on a Z_n qui prend ses valeurs dans $[0, n\theta]$. Donc $f_{Z_n}(z) = 0$ si $z \notin [0, n\theta]$ et pour $z \in [0, n\theta]$,

$$\mathbb{P}[Z_n \leq z] = \mathbb{P}[X_n \geq \theta - z/n] = 1 - \mathbb{P}[X_n \leq \theta - z/n] = 1 - \int_0^{\theta - z/n} f_n(x)dx.$$

Donc en dérivant, on a $f_{Z_n}(z) = \frac{1}{n} f_n(\theta - z/n) = (\theta - z/n)^{n-1}\,\theta^{-n} = \frac{1}{\theta}\,(1 - \frac{z}{n\theta})^{n-1}$.
Lorsque $n \to \infty$, pour tout $z \geq 0$, on a $f_{Z_n}(z) = \frac{1}{\theta}\,\exp\Big((n-1)\ln(1 - \frac{z}{n\theta})\Big) \xrightarrow[n\to\infty]{} \frac{1}{\theta}e^{-z/\theta}$ et pour $z < 0$, $f_{Z_n}(z) = 0$. On en déduit donc que $Z_n \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{E}(1/\theta)$.

□

2. **(Sur 22 points)** Soit $\sigma^2 > 0$ un réel inconnu et soit $\varepsilon = (\varepsilon_t)_{t\in\mathbf{Z}}$ un bruit blanc gaussien de variance σ^2. Pour tout $t \in \mathbf{N}$, on définit X_t par

$$X_t = 2X_{t-1} - X_{t-2} + \varepsilon_t,$$

et par hypothèse $X_{-1} = b$ et $X_{-2} = a$, où a et b sont des réels inconnus.

(a) On pose $U_t = X_t - X_{t-1}$ pour $t \in \mathbf{N}$. Montrer que pour tout $t \in \mathbf{N}$, $U_t = (b-a) + \sum_{i=0}^t \varepsilon_i$ **(1 pt)**.

(b) En déduire que pour $t \in \mathbf{N}$, $X_t = b + (t+1)(b-a) + \sum_{i=0}^t (t-i+1)\varepsilon_i$ **(1 pt)**. Quelle est la tendance, la saisonnalité et le bruit de $(X_t)_{t\in\mathbf{N}}$ **(0.5 pts)**? Déterminer la variance et la loi de X_t pour $t \in \mathbf{N}$ **(2 pts)**. $(X_t)_{t\in\mathbf{N}}$ est-elle stationnaire **(0.5 pts)**?

(c) On suppose observée une trajectoire $(X_0, \cdots, X_n)$ et on considère la trajectoire $(U_1, \cdots, U_n)$ obtenue à partir de $(X_0, \cdots, X_n)$. Montrer que pour $i \le j$, $\mathrm{cov}(U_i, U_j) = (i+1)\,\sigma^2$ **(2 pts)**. En déduire la loi du vecteur $\mathbf{U}_n = {}^t(U_1, \cdots, U_n)$ (on notera J_n la matrice de covariance de $\mathbf{U}_n$ que l'on explicitera) **(1.5 pts)**.

(d) On note $\theta = b - a$. Montrer que $\overline{U}_n = \frac{1}{n}(U_1 + \cdots + U_n)$ est un estimateur sans biais de θ **(0.5 pts)**. Quelle est la loi de $\overline{U}_n$ **(2.5 pts)**? Est-ce un estimateur convergent de θ **(1 pt)**?

(e) Vérifier que $J_n^{-1} = \dfrac{1}{\sigma^2}\Sigma_n^{-1}$ avec $\Sigma_n^{-1} = \begin{pmatrix} 3/2 & -1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}$ **(1 pt)**. Montrer que l'estimateur $\widehat{\theta}_n$ par maximum de vraisemblance de θ vérifie $\widehat{\theta}_n = ({}^t\mathbf{1}_n \Sigma_n^{-1}\mathbf{1}_n)^{-1}\,{}^t\mathbf{1}_n\Sigma_n^{-1}\mathbf{U}_n$ avec $\mathbf{1}_n = {}^t(1, \cdots, 1)$ **(2 pts)**. Que vaut $\widehat{\theta}_n$? **(1 pt)** Quelle est sa loi? **(0.5 pts)** Est-ce un estimateur convergent de θ? **(0.5 pts)** Comparer son risque quadratique avec celui de $\overline{U}_n$ et conclure **(0.5 pts)**.

(f) Montrer par récurrence que le déterminant de J_n est égal à $2\,\sigma^{2n}$ **(2 pts)**. Utiliser ce qui précède pour montrer que l'estimateur par maximum de vraisemblance $\widehat{\sigma}_n^2$ de σ^2 est tel que $\widehat{\sigma}_n^2 = \dfrac{1}{n}\,{}^t\big(\mathbf{U}_n - U_1\mathbf{1}_n\big)\Sigma_n^{-1}\big(\mathbf{U}_n - U_1\mathbf{1}_n\big)$ **(2 pts)**.

Proof. (a) On a facilement $U_{t+1} = U_t + \varepsilon_{t+1}$ pour tout $t \in \mathbf{N}$ et $U_0 = X_0 - b = (b-a) + \varepsilon_0$. Par itération, $U_1 = U_0 + \varepsilon_1 = (b-a) + \varepsilon_0 + \varepsilon_1$ et pour tout $t \in \mathbf{N}$, $U_t = (b-a) + \sum_{i=0}^t \varepsilon_i$.

(b) (Il est aussi possible de montrer le résultat par récurrence sur t...). On a $X_t = X_{t-1} + (b-a) + \sum_{i=0}^t \varepsilon_i$ pour tout $t \in \mathbf{N}$. Avec $X_{-1} = b$, on a par itération,

$$X_t = X_{t-2} + (b-a) + \sum_{i=0}^{t-1}\varepsilon_i + (b-a) + \sum_{i=0}^{t}\varepsilon_i = X_{t-2} + 2(b-a) + \sum_{i=0}^{t-1}\varepsilon_i + \sum_{i=0}^{t}\varepsilon_i.$$

En continuant l'itération, on en arrive à ce que $X_t = X_0 + t(b-a) + \sum_{j=0}^{t-1}\sum_{i=0}^{t-j}\varepsilon_i$ pour $t \in \mathbf{N}^*$. D'une part, on a $X_0 = 2b - a + \varepsilon_0$. D'autre part, dans la double somme, pour tout $i \in \{0, 1, \cdots, t\}$, ε_i est présent $t+1-i$ fois, d'où la réécriture de la double somme sous la forme $\sum_{i=0}^t (t-i+1)\varepsilon_i$.
On calcule $\mathbb{E}X_t = b + (t+1)(b-a)$ pour $t \in \mathbf{N}$. Il n'y a pas de fonction périodique dans cette espérance, donc (X_t) n'admet pas de saisonnalité et a pour tendance une fonction affine $t \mapsto b + (t+1)(b-a)$. La partie bruit de (X_t) est $Z_t = \sum_{i=0}^t (t-i+1)\varepsilon_i$.
On a $\mathrm{var}(X_t) = \sigma^2 \sum_{i=0}^t \sum_{j=0}^t (t-i+1)(t-j+1) = \sigma^2 \sum_{i'=1}^{t+1}\sum_{j'=1}^{t+1} i'j' = \frac{\sigma^2}{4}(t+1)^2(t+2)^2$.
Comme (ε_t) est un processus gaussien la partie bruit de X_t est gaussienne comme combinaison linéaire d'un processus gaussien.
Ainsi la loi de X_t est $\mathcal{N}\big(b + (t+1)(b-a)\,,\ \frac{\sigma^2}{4}(t+1)^2(t+2)^2\big)$.
L'espéance et la tendance de (X_n) dépendant du temps, on en déduit que (X_n) n'est pas stationnaire.

(c) On a $U_i = (b-a) + \sum_{i=0}^t \varepsilon_i$, donc $\mathrm{cov}(U_j, U_j) = \sum_{p=0}^i \sum_{q=0}^j \mathrm{cov}(\varepsilon_p, \varepsilon_q)$. Comme les ε_k forment un bruit blanc, ils sont tous centrés et indépendants, donc $\mathrm{cov}(\varepsilon_p, \varepsilon_q) = \sigma^2$ si $p = q$, et $\mathrm{cov}(\varepsilon_p, \varepsilon_q) = 0$ sinon. On en déduit que $\mathrm{cov}(U_j, U_j) = \sum_{p=0}^i \sigma^2 = (i+1)\sigma^2$.
Comme $\mathbf{U}_n$ s'écrit comme une matrice multipliée par le vecteur $(\varepsilon_0, \cdots, \varepsilon_n)$ qui est gaussien, on en déduit que $\mathbf{U}_n$ est gaussien. L'espérance de $\mathbf{U}_n$ est $(b-a)\mathbf{1}_n$ avec $\mathbf{1}_n = {}^t(1, \cdots, 1)$. Sa matrice de covariance J_n s'écrit:

$$J_n = \sigma^2 \begin{pmatrix} 2 & 2 & 2 & \cdots & 2 \\ 2 & 3 & 3 & \cdots & 3 \\ 2 & 3 & 4 & \cdots & 4 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 2 & 3 & 4 & \cdots & (n+1) \end{pmatrix}$$

On en déduit donc que $\mathbf{U}_n \overset{\mathcal{L}}{\sim} \mathcal{N}\big((b-a)\mathbf{1}_n\,,\ J_n\big)$.

(d) Il est clair que comme $\mathbb{E}U_i = \theta$ pour tout $i = 1, \cdots, n$, alors $\mathbb{E}\overline{U}_n = \theta$.
On peut écrire que $\overline{U}_n = \frac{1}{n}\,{}^t\mathbf{1}_n\mathbf{U}_n$. Donc $\mathrm{var}(\overline{U}_n) = \frac{\sigma^2}{n^2}\,{}^t\mathbf{1}_n J_n \mathbf{1}_n$. On peut aussi écrire, grâce à la formule de la covariance détérminée ci-dessus, $\mathrm{var}(\overline{U}_n) = \frac{1}{n^2}\Big(\sum_{i=1}^n \mathrm{cov}(U_i, U_i) + 2\sum_{i=1}^n\sum_{j=1}^{i-1}\mathrm{cov}(U_i, U_j)\Big) = \frac{\sigma^2}{n^2}\Big(\sum_{i=1}^n (i+1) + 2\sum_{i=1}^n\sum_{j=1}^{i-1}(j+1)\Big) = \frac{\sigma^2}{2n^2}\Big(n(n+3) + 2\sum_{i=1}^n (i-1)(i+2)\Big) = \frac{\sigma^2}{6n}\big(2n^2 + 9n + 1\big)$.
Comme $\overline{U}_n$ est aussi une variable gaussienne (combinaison linéaire d'un vecteur gaussien), sa variance ne tendant pas vers 0, $\overline{U}_n$ ne converge pas vers θ.

(e) On vérifie l'expression de Σ_n^{-1} en effectuant la multiplication matricielle de Σ_n^{-1} et J_n.
Pour trouver l'estimateur par maximum de vraisemblance de θ, il convient de maximiser la log-vraisemblance, ce qui revient à minimiser ${}^t(\mathbf{U}_n - \theta\mathbf{1}_n)\Sigma_n^{-1}(\mathbf{U}_n - \theta\mathbf{1}_n)$. En dérivant par rapport à θ, on obtient l'expression demandée. De plus, comme on a une fonction quadratique en θ qui tend vers l'infini quand $|\theta|$ tend vers l'infini, on a bien obtenu un mimimum.
En calculant $\widehat{\theta}_n$, on obtient $\widehat{\theta}_n = U_1$. Sa loi est donc une loi gaussienne centrée en θ et de variance $2\sigma^2$. Ce n'est pas un estimateur convergent de θ. Son risque quadratique, qui est $2\sigma^2$ est cependant plus petit que celui de la moyenne empirique $\overline{U}_n$: c'est donc un meilleur estimateur de θ.

(f) On montre que le déterminant vaut $2\sigma^{2n}$ en transformant J_n en effectuant des soustractions de la ligne i par la ligne $i+1$, et en développant par rapport à la première colonne finale où il y a un 2.
La log-vraisemblance vaut donc: $-n\ln(2\pi) - n\ln(\sigma^2) - \ln(2)/2 - \frac{1}{2\sigma^2}{}^t(\mathbf{U}_n - \theta\mathbf{1}_n)\Sigma_n^{-1}(\mathbf{U}_n - \theta\mathbf{1}_n)$. En dérivant par rapport à σ^2, on arrive à l'expression de $\widehat{\sigma}^2$.

$\square$